Proc. Natl. Acad. Sci. USA
Vol. 88, pp. 921–925, February 1991
Biochemistry

CGG: An unassigned or nonsense codon in *Mycoplasma capricolum*

(cell-free translation/genetic code/termination codon/AT pressure/codon capture)

TAKANORI OBA, YOSHIKI ANDACHI, AKIRA MUTO, AND SYOZO OSAWA

Laboratory of Molecular Genetics, Department of Biology, School of Science, Nagoya University, Chikusa-Ku, Nagoya 464-01, Japan

Communicated by Motoo Kimura, November 1, 1990

ABSTRACT **CGG is an arginine codon in the universal genetic code. We previously reported that in *Mycoplasma capricolum*, a relative of Gram-positive eubacteria, codon CGG did not appear in coding frames, including termination sites, and $tRNA^{Arg}_{CCG}$ pairing with codon CGG, was not detected. These facts suggest that CGG is a nonsense (unassigned and untranslatable) codon—i.e., not assigned to arginine or to any other amino acid. We have investigated whether CGG is really an unassigned codon by using a cell-free translation system prepared from *M. capricolum*. Translation of synthetic mRNA containing in-frame CGG codons does not result in "read-through" to codons beyond the CGG codons—i.e., translation ceases just before CGG. Sucrose-gradient centrifugation profiles of the reaction mixture have shown that the bulk of peptide that has been synthesized is attached to 70S ribosomes and is released upon further incubation with puromycin. The result suggests that the peptide is in the P site of ribosome in the form of peptidyl-tRNA, leaving the A site empty. When in-frame CGG codons are replaced by UAA codons in mRNA, no read-through occurs beyond UAA, just as in the case of CGG. However, the synthesized peptide is released from 70S ribosomes, presumably by release factor 1. These data suggest strongly that CGG is an unassigned codon and differs from UAA in that CGG is not used for termination.**

We have reported that in *Mycoplasma capricolum*, a Gram-positive relative of eubacteria that lacks a cell wall, $tRNA^{Arg}_{CCG}$ (as well as the gene for this tRNA), which is needed to read the CGG codon in eubacteria, is absent (1, 2) and that no CGG codons have been found in a total of 6814 codons in *M. capricolum* genes examined (1, 3). These results suggest that CGG may be an unassigned (nonsense) codon. To test this, we have constructed a cell-free system of *M. capricolum*, so as to translate added synthetic mRNA containing in-frame CGG codons. The results show that CGG codon is not translated and is not used as a termination codon.

MATERIALS AND METHODS

Preparation of the S-30 Fraction. *M. capricolum* [American Type Culture Collection 27343 (kid)] cells were grown at 37°C, without shaking, in a medium containing 2.2% PPLO broth (Difco), 0.1% yeast extract (Difco), 0.2% glucose, 20 mM Tris·HCl (pH 7.6), 0.01% thallous acetate, penicillin G at 200 units/ml, and 1% calf serum (GIBCO) (4). Cells were quickly chilled at middle-logarithmic phase and collected by centrifugation at 7500 × *g* for 20 min at 4°C. The collected cells were immediately ground with quartz sand and suspended in buffer A (10 mM Tris·HCl, pH 7.8/10 mM magnesium acetate/60 mM NH_4Cl/6 mM 2-mercaptoethanol) containing DNase I at 2 μg/ml. The extract was centrifuged at 22,000 × *g* for 30 min to remove cell debris and then at 33,000 × *g* for 30 min to obtain the S-30 fraction. The S-30 fraction so prepared was incubated at 37°C for 30 min in translation buffer (see below) with 0.5 μM each of 20 amino acids to reduce preexisting mRNA. The fraction was then dialyzed at 4°C overnight against buffer A and stored at −70°C before use. Final protein concentration was 19 mg/ml.

Preparation of Synthetic mRNA. DNA fragments containing the reading frame corresponding to mRNA to be synthesized (see below) and restriction enzyme sites for cloning were synthesized by Gene Assembler Plus (Pharmacia LKB), ligated to plasmid vector Bluescript II KS+ (Stratagene) at the *Pst* I and *Sac* I sites, and transfected to *Escherichia coli* JM109 cells. The plasmids containing the correct insert were selected by DNA sequencing, and the DNA was prepared from the plasmids. mRNA was transcribed from the DNA with T7 RNA polymerase (5) and stored at −70°C before use.

Cell-Free Translation. The standard reaction mixture contained, in a total of 100 μl, 5 μmol of Tris·HCl (pH 7.8), 0.35 μmol of magnesium acetate, 6 μmol of NH_4Cl, 0.1 μmol of dithiothreitol, 50 pmol each of methionine, isoleucine, arginine, and tyrosine, 20 μCi (1 Ci = 37 GBq) of each one of ^{3}H-labeled amino acids, 10 μl of the S-30 fraction, and various amounts of synthetic mRNA. The reaction mixture was incubated at 37°C for 0–30 min. Aliquots of the mixture were heated in 10% trichloroacetic acid for 20 min at 90°C. Hot trichloroacetic acid-insoluble material was collected on nitrocellulose filters, and the radioactivity was measured by a scintillation counter (Tri-Carb).

Sucrose-Gradient Centrifugation. The reaction mixture (500 μl) containing 100 μg of synthetic mRNA was incubated at 37°C for 20 min, layered on 25 ml of 5–20% (wt/vol) sucrose gradient made in the translation buffer (50 mM Tris·HCl/3.5 mM magnesium acetate/60 mM NH_4Cl/0.6 mM 2-mercaptoethanol), and centrifuged in a Beckman SW 25.1 rotor at 35,000 × *g* for 13 hr at 1°C. Fractions were collected from the bottom of the tube, and hot trichloroacetic acid-insoluble radioactivity of each fraction was measured as described above.

Materials. [^{3}H]Isoleucine (92 Ci/mmol), [^{3}H]arginine (58 Ci/mmol), [^{3}H]tyrosine (50 Ci/mmol), and [^{32}P]pCp (3000 Ci/mmol) were purchased from Amersham. Other reagents were from Wako Pure Chemical, Osaka.

RESULTS

The synthetic mRNAs used in this study are shown in Fig. 1. These contained a pair of each one of the six universal arginine codons, CGU, CGC, CGA, CGG, AGA, and AGG, or a UAA termination codon in the reading frame. The codons to be tested (test codons) were preceded by 10 isoleucine codons (7 AUC and 3 AUU) and followed by 5 tyrosine UAC codons and one UAA termination codon. Isoleucine codons were used to find whether mRNA was translated 5′ to the test codon. Incorporation of [^{3}H]tyrosine was the criterion for the efficiency of read-through of the test codon.

These synthetic mRNAs were translated by using [^{3}H]tyrosine as the labeled amino acid (Fig. 2). All mRNAs con-

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

pppGGGCGAAUUGGAGCUC AGGAGG GUUAC
<S.D.>

AUG AUCAUCAUCAUCAUCAUCAUCAUUAUUAUU CGGCGG UACUACUACUACUAC UAA
Met <----------10 x Ile----------> <---5 x Tyr---> Stop

CUGCAGGAAUUCGAUAUCAAGCU$_{OH}$

FIG. 1. Sequences of synthetic mRNAs. The synthetic mRNA used in this study includes, from 5′ to 3′ end, the Shine–Dalgarno (S.D.) sequence, an initiation codon AUG, 10 AUY isoleucine codons, 2 CGG codons, 5 UAC tyrosine codons, and a termination codon UAA, followed by 23 nucleotides. The other mRNAs with the same sequences, but with CGG replaced by the other universal arginine codons (CGU, CGC, CGA, AGA, or AGG) or by termination codon UAA, were also prepared and used for controls and comparisons.

taining universal arginine codons except CGG incorporated [^{3}H]tyrosine into hot-trichloroacetic acid-insoluble material, whereas no incorporation was seen with mRNA containing CGG codons [mRNA(CGG)]. The same result was obtained with mRNA containing UAA codons at these sites (Fig. 1). These data suggest that translation of mRNA(CGG) or mRNA(UAA) stopped at some point before tyrosine codons were reached.

To determine the codon at which translation was blocked on mRNA, the incorporation of each one of three ^{3}H-labeled amino acids, isoleucine, arginine, and tyrosine, was measured for mRNA(CGA), mRNA(CGG), and mRNA(UAA). Appreciable incorporation of each of these three amino acids took place with mRNA(CGA) (Fig. 3*A*). In contrast, no or only a slight incorporation of [^{3}H]arginine was observed either with mRNA(UAA) or mRNA(CGG), whereas significant incorporations of [^{3}H]isoleucine were seen with both these mRNAs (Fig. 3 *B* and *C*). These data indicate that, although CGA codons were translated as arginine with mRNA(CGA), translation did not continue or scarcely continued beyond UAA with mRNA(UAA) or CGG with mRNA(CGG), although a small amount of [^{3}H]arginine was incorporated into hot trichloroacetic acid-insoluble material when mRNA(CGG) was used (see below).

To investigate the state of isoleucine peptides synthesized with mRNA(CGG), the reaction mixture that had been incubated with [^{3}H]isoleucine was centrifuged on a sucrose gradient. The bulk of the incorporated [^{3}H]isoleucine was detected in the 70S ribosome fraction with some in the soluble fraction (top of the gradient) (Fig. 4*A*). When a similar experiment was done with mRNA(UAA), the incorporated

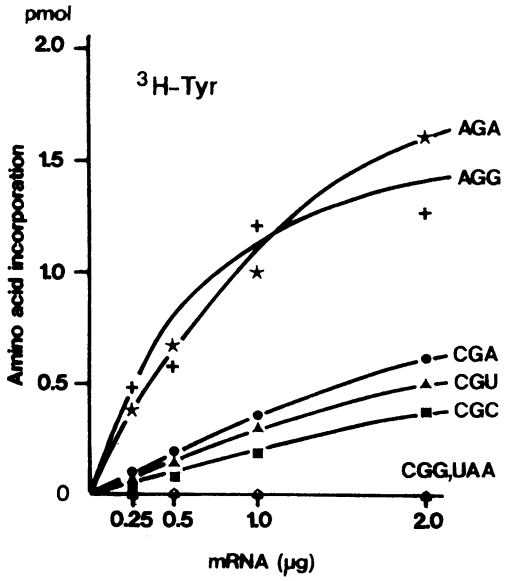

FIG. 2. Incorporation of [^{3}H]tyrosine into hot trichloroacetic acid-insoluble material in cell-free translation of various synthetic mRNAs. The reaction mixture (10 μl) containing 0–2 μg of the synthetic mRNA was incubated at 37°C for 30 min. Incorporation of [^{3}H]tyrosine into hot trichloroacetic acid-insoluble material was measured. In-frame codons in the synthetic mRNA are shown by triplets.

[^{3}H]isoleucine was recovered almost exclusively in the soluble fraction (Fig. 4*B*). These experiments suggest that synthesized peptide containing isoleucine was released from ribosomes at the UAA termination codon when mRNA(UAA) was used, whereas with mRNA(CGG) the peptide could not be released at the CGG codon. In short, codon CGG blocked translation but could not release the synthesized peptide. Addition of puromycin led to the release, from 70S ribosomes, of the [^{3}H]isoleucine-labeled peptides formed with mRNA(CGG), so that these were recovered in the soluble fraction (Fig. 4*C*). Puromycin mimics an amino acid linked to the 3′ end of tRNA, so that it enters the free A site of the ribosome and releases the peptide from peptidyl-tRNA on the P site. It was concluded that [^{3}H]isoleucine-peptidyl-tRNAs on 70S ribosomes stayed at the P sites, and the A site was almost free from aminoacyl-tRNA (see Fig. 7*B*).

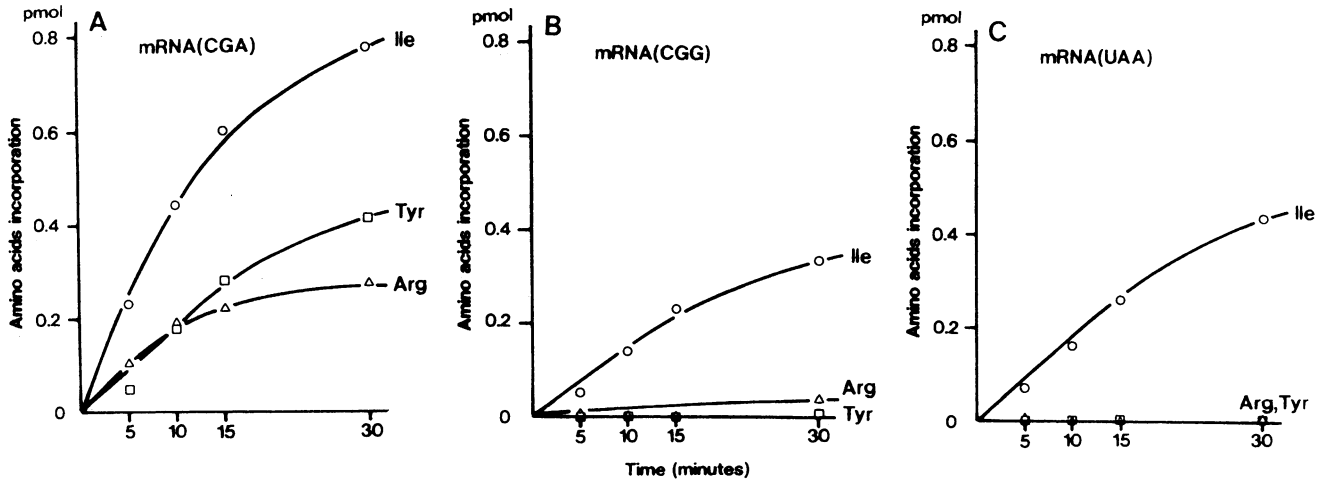

FIG. 3. Incorporation of ^{3}H-labeled amino acids into hot trichloroacetic acid-insoluble material in cell-free translation of synthetic mRNA containing codons CGA (*A*), CGG (*B*), and UAA (*C*), respectively. The reaction mixture (100 μl) containing [^{3}H]isoleucine, [^{3}H]tyrosine, or [^{3}H]arginine, and 25 μg of mRNA was incubated at 37°C. Aliquots (10 μl) were taken from the reaction mixture at time intervals, and hot trichloroacetic acid-insoluble radioactivity was measured.

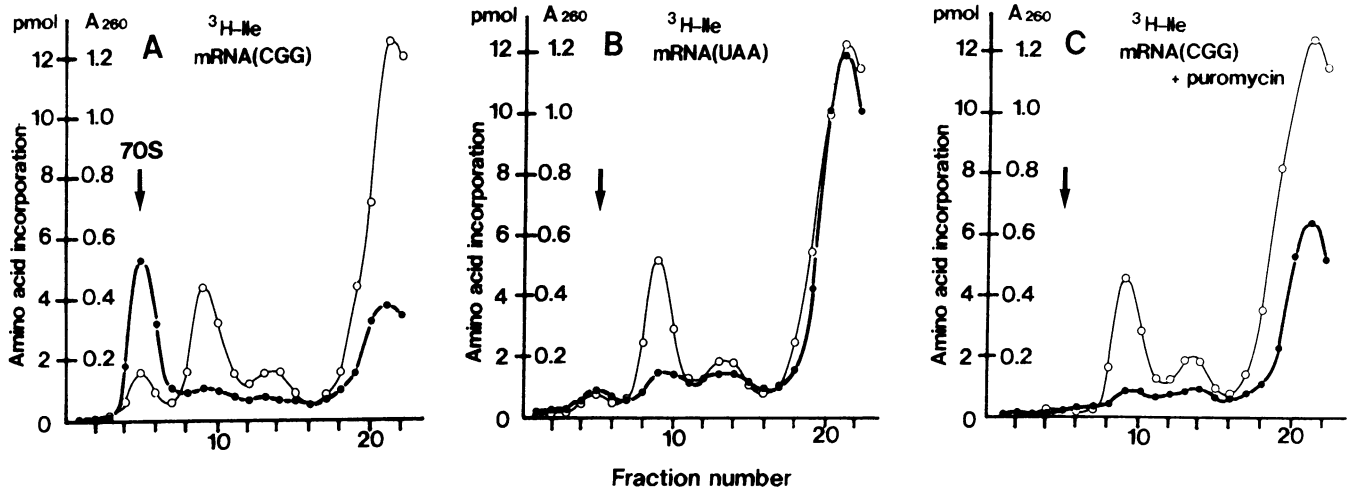

FIG. 4. Sucrose-gradient centrifugation of reaction mixture labeled with [^{3}H]isoleucine. The reaction mixture (500 μl) containing 100 μg of synthetic mRNA with codons CGG (*A* and *C*) or UAA (*B*) and [^{3}H]isoleucine was incubated at 37°C for 20 min. Mixture *C* was further incubated with 1 mM puromycin at 37°C for 10 min. The mixture was centrifuged on 5–20% sucrose gradient and fractionated into 25 tubes. UV absorbance at 260 nm (○) and hot trichloroacetic acid-insoluble radioactivity (●) of each fraction were measured.

As noted above, a small amount of [^{3}H]arginine was incorporated into hot trichloroacetic acid-insoluble material recovered from mRNA(CGG). The [^{3}H]arginine-containing hot trichloroacetic acid-insoluble material was found exclusively in the soluble fraction (Fig. 5). It is probable that some occasional additions of arginine to the C terminus of the peptide at the CGG site through incorrect pairing of CGG with arginyl-tRNA led to a release of the peptidyl-tRNA^{Arg} from ribosomes. If so, a small amount of trichloroacetic acid-insoluble [^{3}H]isoleucine radioactivity in the soluble fraction observed with mRNA(CGG) (see Fig. 4*A*) would represent the released peptidyl-tRNA. To identify the nature of the tRNA involved, the [^{3}H]isoleucine (or arginine)-labeled soluble fraction after translation of mRNA(CGG), or mRNA(UAA) as control, was chromatographed on a BD-cellulose column. Free [^{3}H]isoleucine (or arginine) was not adsorbed on the column. [^{3}H]Isoleucinyl (or arginyl)-tRNA was eluted with a buffer containing 1.8 M NaCl/30% (vol/vol) ethanol. ^{3}H-labeled peptidyl-tRNA or peptides were not eluted. The column content was then incubated with 1 M NaCl/1 M Tris·HCl buffer, pH 9.0, at 37°C for 60 min to cleave the ester bond between tRNA, if any, and the peptide. The liberated RNA fraction was examined by PAGE. Three RNA bands were observed (Fig. 6, lanes a and b). The upper two bands were identified as 4.5S RNA, and the lowest one was identified as $tRNA^{Arg}_{ICG}$ by T1 mapping. From this, we conclude that the material liberated from ribosomes with mRNA(CGG) was peptidyl-$tRNA^{Arg}_{ICG}$. In the control with mRNA(UAA), only 4.5S bands were seen without the tRNA band (Fig. 6, lane c), indicating that peptide synthesis was terminated at codon UAA, followed by liberation of the peptides (but not peptidyl-tRNA) from ribosomes.

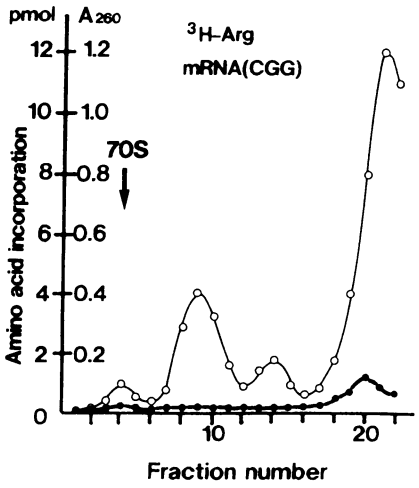

FIG. 5. Sucrose-gradient centrifugation of reaction mixture labeled with [^{3}H]arginine. The same as for Fig. 4*A* but labeled with [^{3}H]arginine.

DISCUSSION

We have shown that CGG, an arginine codon in the universal genetic code, cannot be translated in the cell-free translation system of *M. capricolum*. Translation ceases before or when codon CGG reaches the A site and the peptidyl-tRNA stays at the P site of ribosomes (Fig. 7*B*). Translation also ceases at codon UAA, but the peptide synthesized is released from ribosomes. It is thus evident that codon CGG cannot be translated as an amino acid codon in the cell-free system of *M. capricolum*. From these data together with the probable absence of codon CGG (1, 3) and of the corresponding $tRNA^{Arg}_{CCG}$ in this bacterium (1), we propose that CGG is an unassigned (nonsense) codon in *M. capricolum*. Termination (stop) codons (UAA and UAG; UGA is a tryptophan codon in *M. capricolum*; see ref. 7) are often referred to as nonsense codons. The termination codons are, however, not nonsensical, and they function to release the synthesized polypeptides from ribosomes by interacting with release factor (RF). On the other hand, "true" nonsense (unassigned) codons, such as CGG in *M. capricolum*, cannot bring about release of the polypeptides from ribosomes (Fig. 7*B*), presumably by lack of the ability to interact with release factor. Occurrence of an unassigned codon CGG implies that the codon table for this bacterium has 63 codons, including termination codons, instead of the 64 in the universal codon table. We define the unassigned or nonsense codon as the codon that does not exist in an organism (or organelle) and cannot make its appearance from mutations of other codons due to lack of the relevant anticodon or release factor corresponding to the codon. Occurrence of unassigned codons seems not to be restricted to *M. capricolum*. We have found that several codons ending with adenosine (ref. 8; A. Kano, T. Ohama, Y.A., A.M., and S.O., unpublished work) have not been detected in the high-G+C bacterium, *Micrococcus luteus*.

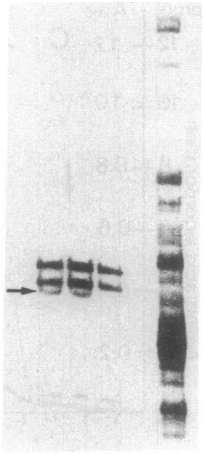

FIG. 6. Identification of tRNA derived from peptidyl-tRNA in the soluble fraction. The reaction mixture (500 μl) containing 100 μg of mRNA with codon CGG (lanes a and b) or UAA (lane c) and [^{3}H]isoleucine (lanes a and c) or [^{3}H]arginine (lane b) was incubated at 37°C for 30 min and centrifuged on sucrose gradient as described for Figs. 4 and 5. The soluble (top) fractions were loaded on a benzoylated DEAE-cellulose (BD-cellulose, Serva) column (1 × 10 cm) equilibrated with buffer (20 mM Tris·HCl, pH 7.8/10 mM $MgCl_2$) containing 0.3 M NaCl, followed by washing successively with the buffer containing 1 M NaCl, 1 M NaCl/4.7% (vol/vol) ethanol, and 1.8 M NaCl/30% ethanol. The column content was then incubated with 1 ml of 1 M Tris·HCl, pH 9.0/1 M NaCl at 37°C for 60 min to liberate tRNAs from peptidyl-tRNAs. The mixture was centrifuged to remove resin. Supernatant containing liberated RNAs was concentrated with Centricon 10 (Amicon), and RNAs were precipitated with ethanol. The 3′ ends of RNAs were labeled with [^{32}P]pCp (cytidine 3′,5′-bisphosphate) as described (6), and subjected to 12% polyacrylamide/7 M urea gel electrophoresis. Labeled RNA bands were detected by autoradiography (lanes a, b, and c). Lane d shows the total "tRNA" fraction from *M. capricolum*. Identification of these RNAs was done as follows: RNAs were eluted from the respective bands with 0.5 M ammonium acetate/10 mM magnesium acetate/0.1% SDS and precipitated with ethanol. RNAs were then partially digested by 0.02 units of RNase T1 (Sankyo) with 3 μg of carrier tRNA in 10 μl of buffer (20 mM sodium citrate, pH 5.0/1 mM EDTA/5.8 M urea) at 55°C for 20 min and subjected to 12% polyacrylamide/7 M urea gel electrophoresis (T1 mapping). Autoradiographic patterns of the T1 mapping were compared with previously identified RNAs (1). Upper two bands are 4.5S RNA (chain length of uppermost RNA is one nucleotide longer than that of the second one); band indicated by an arrow is $tRNA^{Arg}_{ICG}$.

There is another codon, CUC leucine, that has not been found in the genes of *M. capricolum*. However, $tRNA^{Leu}_{UAG}$, in which the first anticodon nucleoside uridine is unmodified, is present in *M. capricolum* (1), and this tRNA can translate all the codons in the CUN family box (N represents U, C, A, or G) including CUC by four-way wobbling. Therefore, codon CUC cannot become an unassigned codon.

A small amount of [^{3}H]arginine was incorporated with mRNA(CGG) without further incorporation of [^{3}H]tyrosine, indicating that elongation of most of the peptides stopped before CGG codon was reached and that there was some "leaked" arginine incorporation by CGG. Because of the lack of $tRNA^{Arg}_{CCG}$ pairing with codon CGG, the A site of ribosome (CGG site) is empty, and the P site is occupied by isoleucine-peptidyl-tRNA (Fig. 7*B*). As a result, transpeptidation would not occur. It is, however, possible that arginyl-$tRNA^{Arg}_{ICG}$ accidentally enters the A site by partial pairing of ICG with CGG. Then the peptide of the isoleucine-peptidyl-tRNA on the P site was transferred to arginyl-tRNA on the A site (Fig. 7*C*). Because inosine at the first position of anticodon cannot pair with guanosine at the third position of the codon according to the wobble rules of Crick (9), pairing of codon CGG with anticodon ICG would not occur or, at best, would be unstable; thus, the peptidyl-$tRNA^{Arg}_{ICG}$ is released from ribosomes after transpeptidation, just as for peptidyl-puromycin (Fig. 7*D*). Indeed, released material has been identified as peptidyl-$tRNA^{Arg}_{ICG}$ (Fig. 6).

According to the "two-out-of-three" hypothesis, $tRNA^{Arg}_{ICG}$ can read codon CGG because the first nucleoside inosine of anticodon ICG and the third nucleoside guanosine of codon CGG do not need to be involved in the pairing (10). In our cell-free system, some read-through of CGG codons occurs when magnesium concentration is raised above 4 mM. At the same time, UAA codons are also read through. Thus, we interpret that the apparent two-out-of-three type reading in this case would be nothing but a misreading of CGG (and UAA) in the cell-free system at high magnesium concentration.

How is the unassignment of the codon produced and what is the evolutionary significance of this phenomenon? *M. capricolum* is characterized by the highest genomic A+T content among prokaryotes (11). Such high A+T has been suggested to result from AT-biased directional mutation pressure (AT pressure) acting on the whole genome (12). A striking effect of AT pressure can be seen in codon usage in *M. capricolum* genes. More than 90% of the third positions of codons are occupied by uridine or adenosine, suggesting that many NNC and NNG codons have been converted to the synonymous NNU and NNA codons by AT pressure. In the arginine family box, codons CGU, CGC, and CGA are translated by a single anticodon ICG by wobble pairing, and codon CGG is translated by anticodon CCG in eubacteria. If CGG codons are all converted to other arginine codons—i.e., to CGU, CGA, and AGR by silent mutations under strong AT pressure, anticodon CCG probably becomes unnecessary

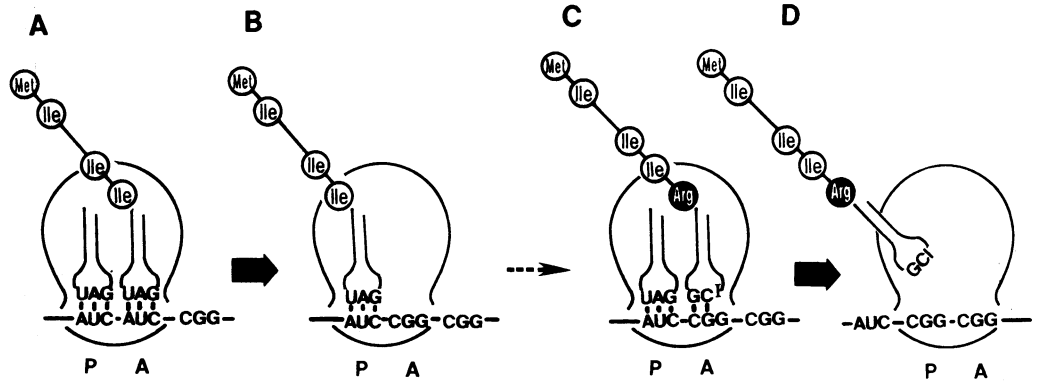

FIG. 7. Models for translation blockage by unassigned codon CGG (*A* to *B*) and occasional false pairing of $tRNA^{Arg}_{ICG}$ with CGG codon (*C*), followed by release of peptidyl-tRNA (*D*).

and disappears. Thus, CGG would become an unassigned codon that has disappeared from coding sequences.

We assume that the production of an unassigned codon is an intermediary step of the codon reassignment during evolution of the genetic code. Our codon-capture theory for the changes in the code involves temporary disappearance of a codon from mRNA by converting to another synonymous codon and of the corresponding tRNA that translates the codon, or a change in the codon–anticodon pairing of the tRNA (producing an unassigned codon), and later reappearance of the codon and tRNA with sometimes the same assignment or at other times a different assignment (13). An example of such a reassignment has been found in *M. capricolum*, where UGA, a universal stop codon, would have become unassigned under AT pressure, converting to UAA with probable deletion of release factor 2, followed by reassignment with (captured by) tryptophan upon appearance of $tRNA^{Trp}_{UCA}$ pairing with UGA (7, 14, 15). In fact, codon UGA is translated as tryptophan as efficiently as UGG is in the *M. capricolum* cell-free system, whereas translation of UGG, but not of UGA, was seen in a similar system of *E. coli* (unpublished data). As in the similar case of UGA, it is highly probable that the unassigned codon CGG in *M. capricolum* will eventually be reassigned to arginine again or to other amino acids, depending on the aminoacylation properties of the newly appearing tRNA with anticodon CCG. Note that the process of codon reassignment, including production of unassigned codons, proceeds under mutation pressures without changing the amino acid sequence of proteins, in accordance with the neutral theory of molecular evolution (16).

We thank Dr. Thomas H. Jukes for critical reading of the manuscript. This work was supported by grants from the Ministry of Education, Science and Culture, Japan.

1. Andachi, Y., Yamao, F., Muto, A. & Osawa, S. (1989) *J. Mol. Biol.* **209**, 37–54.
2. Muto, A., Andachi, Y., Yuzawa, H., Yamao, F. & Osawa, S. (1990) *Nucleic Acids Res.* **18**, 5037–5043.
3. Ohkubo, S., Muto, A., Kawauchi, Y., Yamao, F. & Osawa, S. (1987) *Mol. Gen. Genet.* **210**, 314–322.
4. Sawada, M., Osawa, S., Kobayashi, H., Hori, H. & Muto, A. (1981) *Mol. Gen. Genet.* **182**, 502–504.
5. Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab., Cold Spring Harbor, NY), 2nd Ed.
6. England, T. E., Bruce, A. G. & Uhlenbeck, O. C. (1980) *Methods Enzymol.* **65**, 65–74.
7. Yamao, F., Muto, A., Kawauchi, Y., Iwami, M., Iwagami, S., Azumi, Y. & Osawa, S. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 2306–2309.
8. Ohama, T., Muto, A. & Osawa, S. (1990) *Nucleic Acids Res.* **18**, 1565–1569.
9. Crick, F. H. C. (1966) *J. Mol. Biol.* **19**, 548–555.
10. Lagerkvist, U. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 1759–1762.
11. Muto, A., Yamao, F. & Osawa, S. (1987) *Prog. Nucleic Acid Res. Mol. Biol.* **34**, 29–58.
12. Muto, A. & Osawa, S. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 166–169.
13. Osawa, S. & Jukes, T. H. (1989) *J. Mol. Evol.* **28**, 271–278.
14. Jukes, T. H. (1985) *J. Mol. Evol.* **22**, 361–362.
15. Osawa, S., Muto, A., Jukes, T. H. & Ohama, T. (1990) *Proc. R. Soc. London Ser. B* **241**, 19–28.
16. Kimura, M. (1983) *The Neutral Theory of Molecular Evolution* (Cambridge Univ. Press, Cambridge, England).